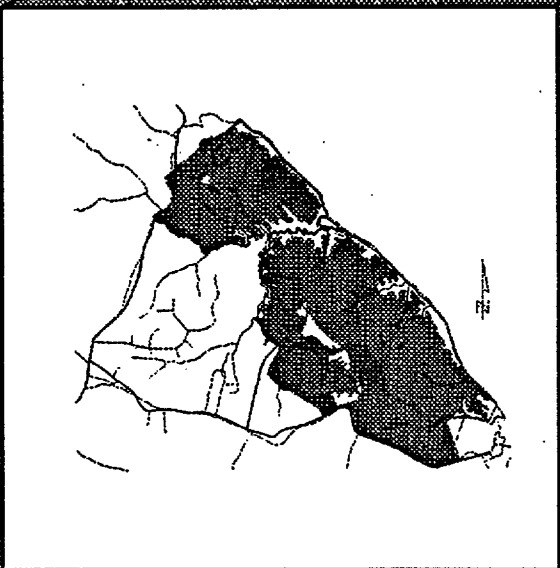
Task 5.3 FY 90
Final Work Product

Geographically-Referencing Natural Resource Inventory Data for York River State Park



Department of Conservation and Recreation
January 1992

Geographically-Referencing Natural Resource Inventory Data for York River State Park

This report was published in part through funds furnished by the Virginia Council on the Environment pursuant to the Coastal Zone Resources Management Grant #NA90AAHCZ796 from the National Oceanic and Atmospheric Administration.

January 1992

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. PURPOSE

The Department of Conservation and Recreation applied for this grant to integrate existing data layers with additional park specific natural resource information (forest type, watershed boundaries, etc.) to demonstrate the applicability of DCR's Arc/Info database as a coastal resource management tool for York River State Park. The development of resource management plans are part of the Division of State Parks' Strategic Plan and the use of geographically referenced data will help meet the goals of this plan.

In addition, a comparison of two methods of data entry was proposed which would evaluate the efficiency of each method. Originally, the grant was to include a feasibility study of view shed analysis on Arc/Info. However, time constraints and initial investigations into this revealed that this could not be covered in the scope of this work.

II. DOCUMENTATION OF DATA LAYERS:

1) Base Map Development

The layers that composed the Base Map were collected from various sources (See Table 1). Road networks, streams and shoreline data layers were available from the Council on the Environment in digital form (taken from USGS Digital Line Graph files). National Wetlands Inventory data was available on Arc/Info through the Department of Conservation and Recreation, Division of Soil and Water Conservation. This information was originally created as an overlay for 7.5-minute USGS topoquads using 1:80000 panchromatic aerial photography. Digital information had been purchased from the USFWS at an earlier date by DGIF and was available on Prime Arc/Info. Park boundaries were taken directly from the Gressitt Topoquad and were already in Arc/Info, while watershed boundaries were initially delineated by the State Water Control Board, based on topographic relief, on the Gressitt Topo.

The watershed boundary was a useful and ecologically valid tool in selecting information to include in the coverage. The boundary was digitized and used to "CLIP" data from another coverage for inclusion. This information was digitized at 1:24000 using a Numonics digitizing tablet purchased as a partial match for this grant.

Table 1 - Base Map Data Layers

Layer	Source	Scale
Roads	USGS DLG Files	1:100000
Streams	USGS DLG Files	1:100000
Shoreline	USGS DLG Files	1:100000
Wetland	USFWS Natl Wetlands Inventory	1:24000
Park Boundary	Gressitt Quadrangle	1:24000
Watershed Boundary	Gressitt Quadrangle	1:24000

Upon completion of the first Base Map on May 13, 1991, it was revealed that the Digital Line Graph File containing the streams information was incomplete. It appears that the smaller scale (1:100000) resulted in the loss of several stream segments. To rectify this, it was decided that the streams should be redigitized from the topoquad. Since this work was being done directly off the topoquad and ephemeral/perennial streams are included in the USGS topoquads, it was decided that the "streams" data layer would be subdivided into two subsets (YRSP/STRMEPH and YRSP/STRMPER). The "streams" coverages were created (YRSP/STRMEPH and YRSP/STRMPER), digitized directly from the topoquad, edited, built, and transformed. The final basemap consisted of the data layers in Table 2 and was completed on June 7, 1991 (Figure 1).

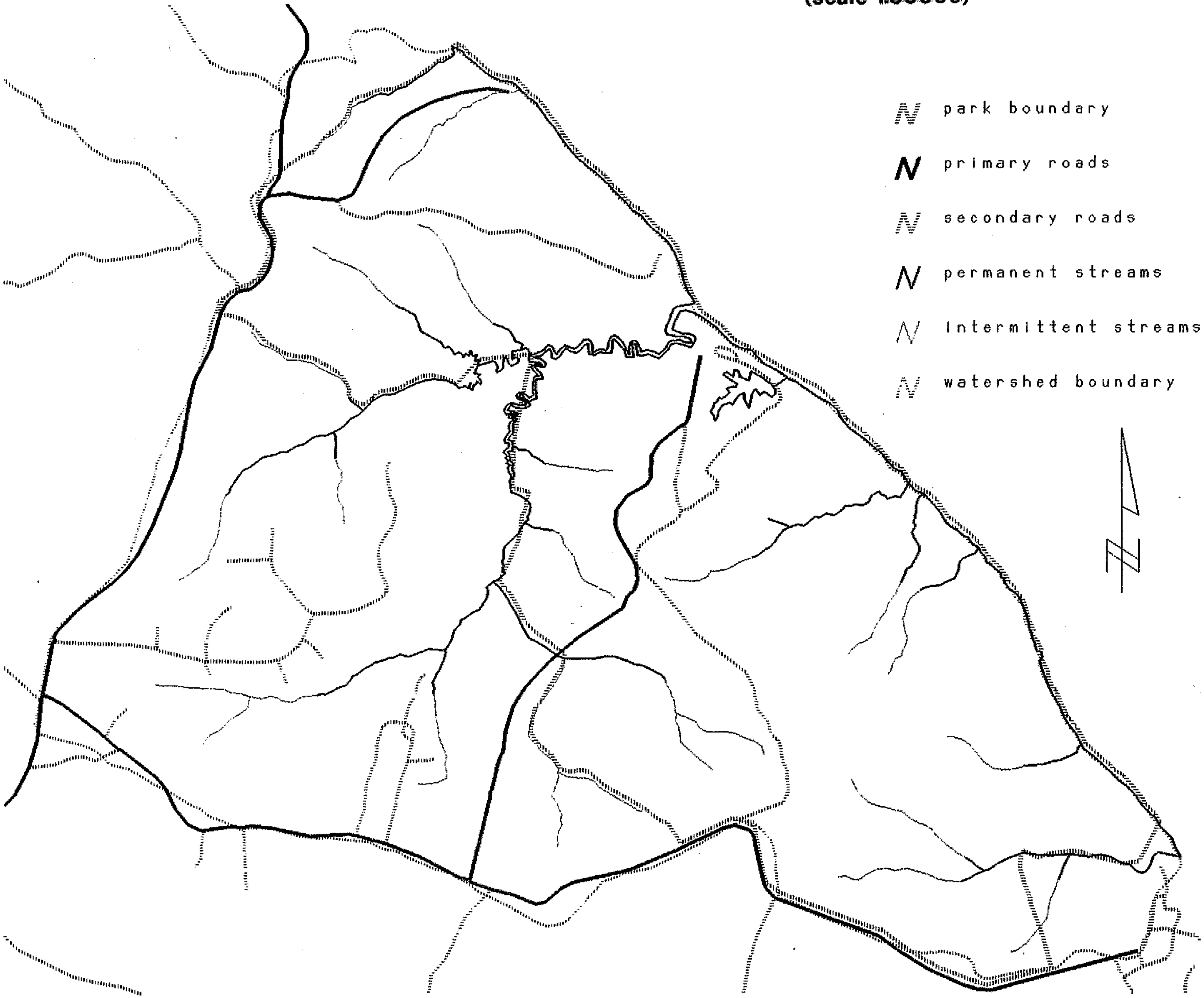
Final coverages and plotfiles are listed in Appendix A. Estimates of time expenditures are listed in Appendix B.

Table 2 - Final Base Map Data Layers

Layer	Source	Scale
Roads	USGS DLG	1:100000
Wetlands	USFWS NWI Maps	1:24000
Streams	USGS Topoquad	1:24000
Ephemeral		
Perennial		
Park Boundary	USGS Topoquad	1:24000
Watershed	USGS Topoquad	1:24000

The "shoreline" data layer was omitted from the final base map because it duplicated much of the NWI coverage (adding no additional information) and in several places, large omissions were made, which make this data layer of little value to site-specific management.

Figure 1: Base Map of York River State Park
(scale 1:30000)



2) Forest Inventory and Database

Forest resource information was available for York River State Park in a crudely-mapped format from a recent forest inventory (Gaston, 1989). Gaston (1989) produced a composite map from two frames of panchromatic aerial photography at a scale of 1:25200 (1" equals 2100'); seventeen forest types were identified using standard photointerpretation methods.

Although the map could be photographically enlarged to approximately the base map scale parallax would most likely result in some distortion of the unrectified image and was believed to make transfer to a standard map scale (i.e., 1:24000) difficult. However, if adequate control points were available with known coordinates, then vectors could be fit by the GIS alleviating the need to locally fit and transfer this data. Control points were taken from road intersections that were identifiable on both the USGS topo-quad and the forest inventory map, entered into Arc/Info and the map was directly digitized. Due to the clustering of most control points in the northwestern portion of the map, the digitized product would not accurately register on the base map (which was anticipated by both Sylvia Terziotti, COE, and Comfort Miller). Therefore, unless evenly distributed control points are available, direct digitization of un-rectified data layers will not result in a usable product.

Ultimately, the forest inventory information was locally-fit and delineated onto the Gressitt topoquad. In addition to problems caused by parallax and edge-matching two aerial photos, forest boundaries with other land uses (i.e., wetlands, open water and open areas) were inconsistent in configuration with known boundaries. In an effort to correct this, a recent (1989) high altitude aerial photo (NAPP) was enlarged to approximately 1:24000, acetate film overlaid and forest/non-forest boundaries were delineated. This overlay was then locally-fit on the topoquad and forest boundaries were digitized into the forest data layer.

The data base design was patterned after the forest inventory datasheet of Gaston (1989) and is presented in Table 3. Forest management coverage is contained in one file (MGMT.UTM) and took approximately four weeks to complete. It is estimated that this can be decreased substantially since several products were digitized prior to selection of a final method. If the final desired product is simply forest cover, and species composition is not important, this estimate could be reduced even further by direct digitization of ground-truthed USGS topoquads (taking approximately 7-10 days to ground-truth and digitize information).

Forest coverage was organized by "TYPE". That is, whether the community consisted of upland hardwoods, bottomland hardwoods, pine and pine-hardwood. Other distinctions were made within the upland hardwood type, but for general management of the park, "upland hardwood" is sufficient. Figure 2 shows an overlay of the forest datalayer (MGMT.UTM) on the basemap with upland hardwood, pine - hardwood mix and bottomland hardwood types selected and shaded.

Figure 2 : Forest Management Datalayer

(scale 1:30000)

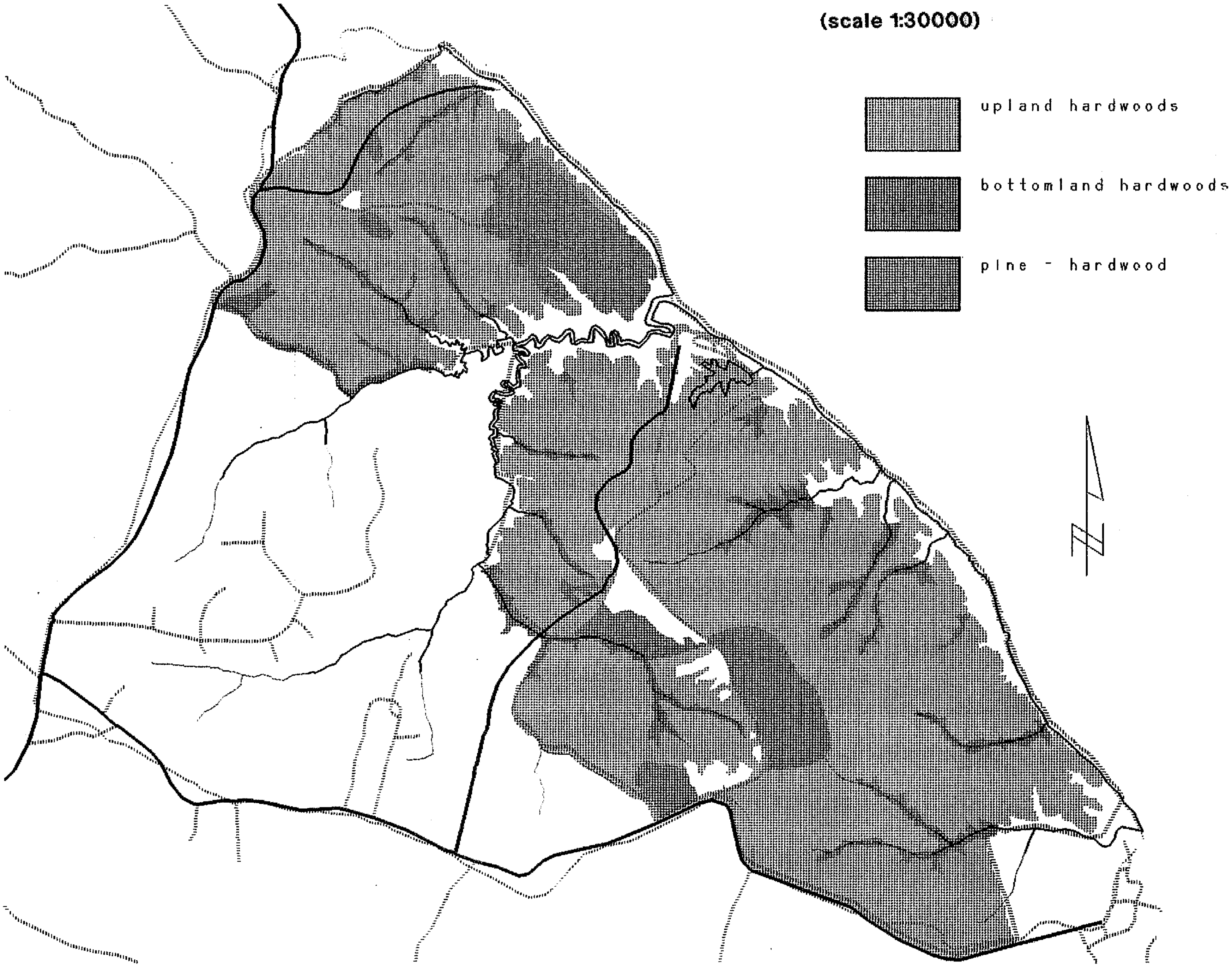


Table 3 - Database Design for Forest Management Data Layer

FOREST MANAGEMENT AT YORK RIVER STATE PARK		
Area #:	-	(2)
Type:	-	(17)
# of Acres:	-	(3)
Species:	-	(40)
	-	(40)
	-	(40)
Age:	-	(10)
Size:	-	(20)
Quality:	-	(30)
Basal Area:	-	(3)
Density (trees/acre):	-	(20)
Growth Rate (inches/yr):	-	(42)
Site Index:	-	(42)

NOTE: The numbers indicate the number of spaces allocated for the responses.

3) Historical and Cultural Sites

There are fifteen sites located in or around York River State Park, based on a survey of existing files at the Department of Historic Resources. UTM coordinates were used to enter data points into Arc/Info, using a keypad. The GENERATE command was used to create the coverage. (Figure 3).

The database for this coverage was developed (Table 4) as a draft and submitted to the DHR for review and comment. DHR is currently revising an existing datasheet and has also informed us that current classifications of archeological data have changed since the mid-70's (when this data was originally collected) and that the reports would need to be revised prior to entering data into Arc/Info. As a result of this, this coverage hasn't been completed.

Figure 3 : Historic and Cultural Sites

(scale 1:24000)

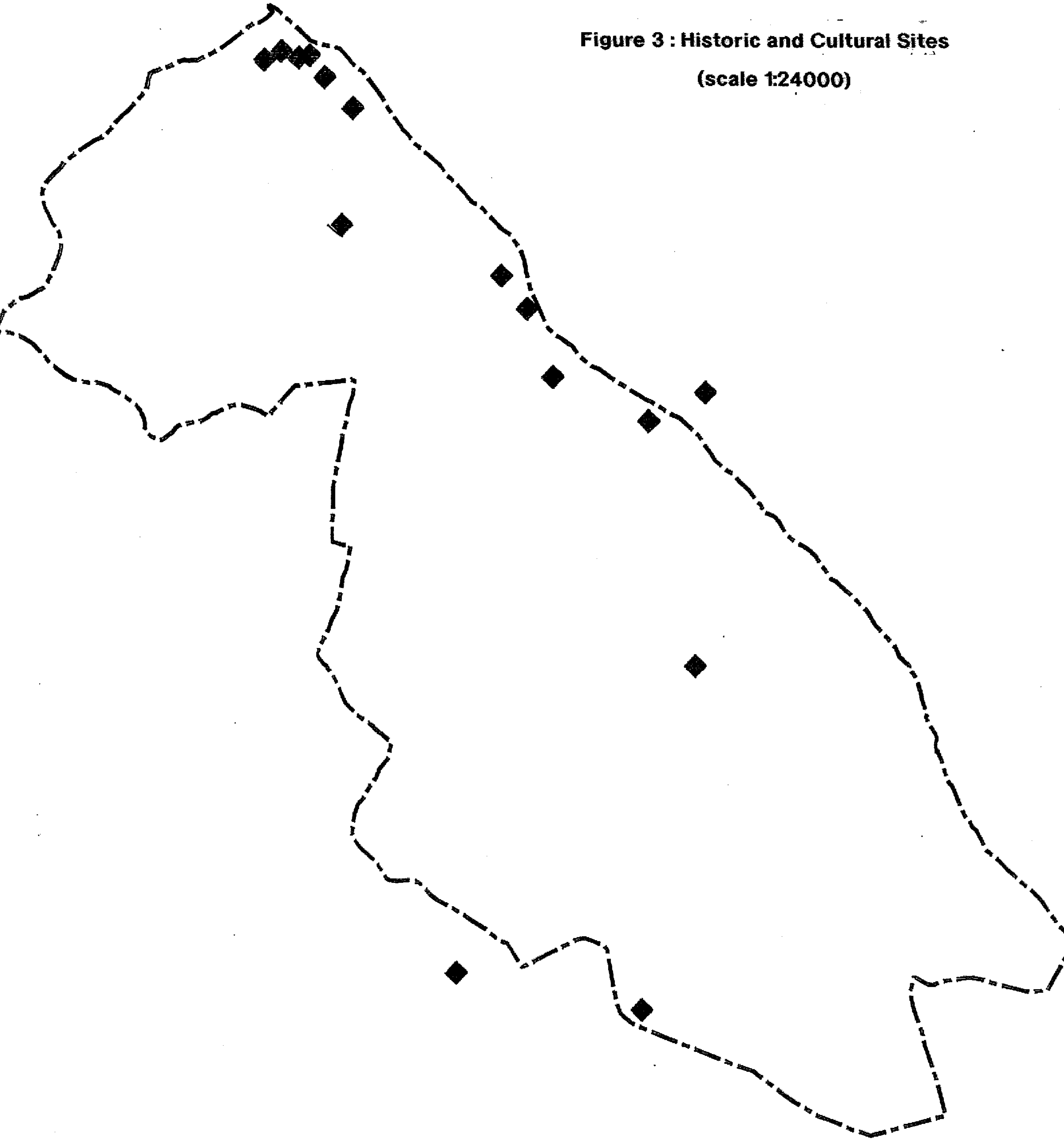


Table 4

HISTORIC INFORMATION YORK RIVER STATE PARK	
Site Number:	__ (10)
Name of Site:	__ (35)
Type of Site:	__ (35)
Cultural Affiliation:	__ (35)
Quad Sheet:	__ (15)
Location:	
UTM Zone:	__ (2)
Easting:	__ (8)
Northing:	__ (9)
Surveyed:	__ (3)
Date Surveyed:	__ (8)
Surveyed by:	__ (35)

4) Land Classification

Land Classification maps were created by the Departmental Land Classification Task Force using a classification system which broadly classifies State lands into three zones (non-sensitive, sensitive and preservation) within a major classification (DCR 1990). Land classifications ultimately drive the development of management units and prescriptions and are derived from environmental, biological and cultural features of each site. Future revisions of these classifications may be refined using a geographic information system, comparing existing classifications with geographically-referenced information to fine tune classifications.

The Land Classification Map (scale = 1:12000) was photo-reduced onto vellum at a scale of 1:24000 and land classification boundaries were transferred to a 7.5-minute USGS topoquad. Boundaries were double-checked for errors and the coverage (LNDCLS) was created, digitized, edited and built. (See Figure 4)

Labels were created for the coverage. In order to assign the same USER_ID to more than one polygon, the SELECT commands were used and the ID's were calculated/applied to the desired values. The data file was created, following the format in Table 5.

Figure 4 : Land Classification at York River State Park

(scale 1:30000)

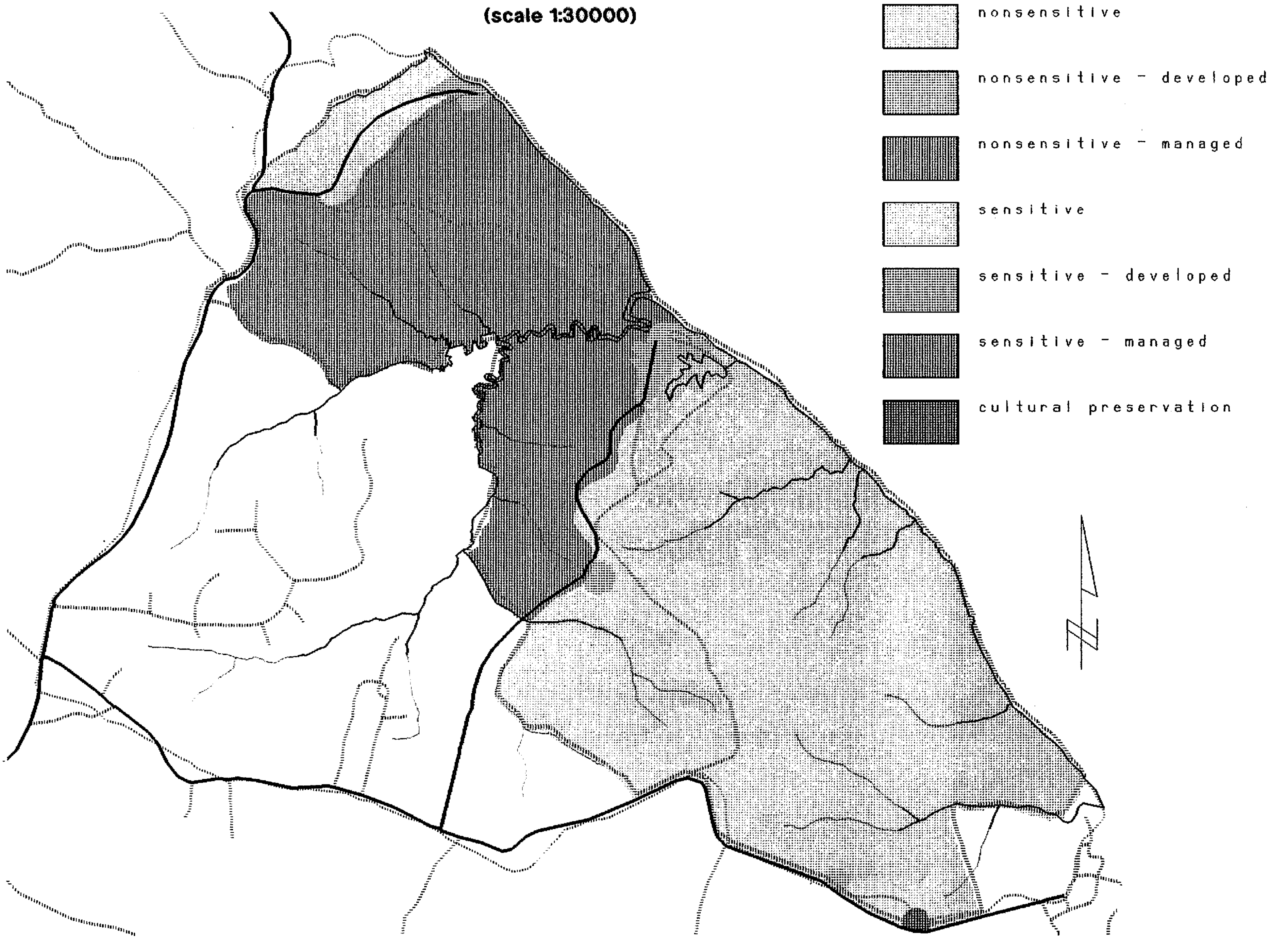


Table 5 - Land Classification Data File

LU-Code	LU-Zone
1	Non-Sensitive
1d	Non-Sensitive (Existing Development)
1m	Non-Sensitive (Management Agreement)
2	Sensitive
2d	Sensitive (Existing Development)
2m	Sensitive (Management Agreement)
3B	Preservation Zone (Cultural)

The final coverage is LNDCLS.CL. (See Appendix A.)

5) Management Units

Management Units were developed as part of the Resource Management Planning process (DCR 1991) on a scale of 1:12000. Photo reductions were made onto vellum and unit boundaries were transferred; coverage (MGTUNIT) was created as described above (see Figure 5).

Based on field data, Natural Heritage staff scientists developed two sets of boundaries for this project: Ecological Boundaries and a larger Conservation Planning Boundary. The Ecological Boundary encompasses the locations of the rare species and the adjacent habitat for them. The Conservation Planning Boundary delineates the larger area within which changes in land use could adversely affect the rare species.

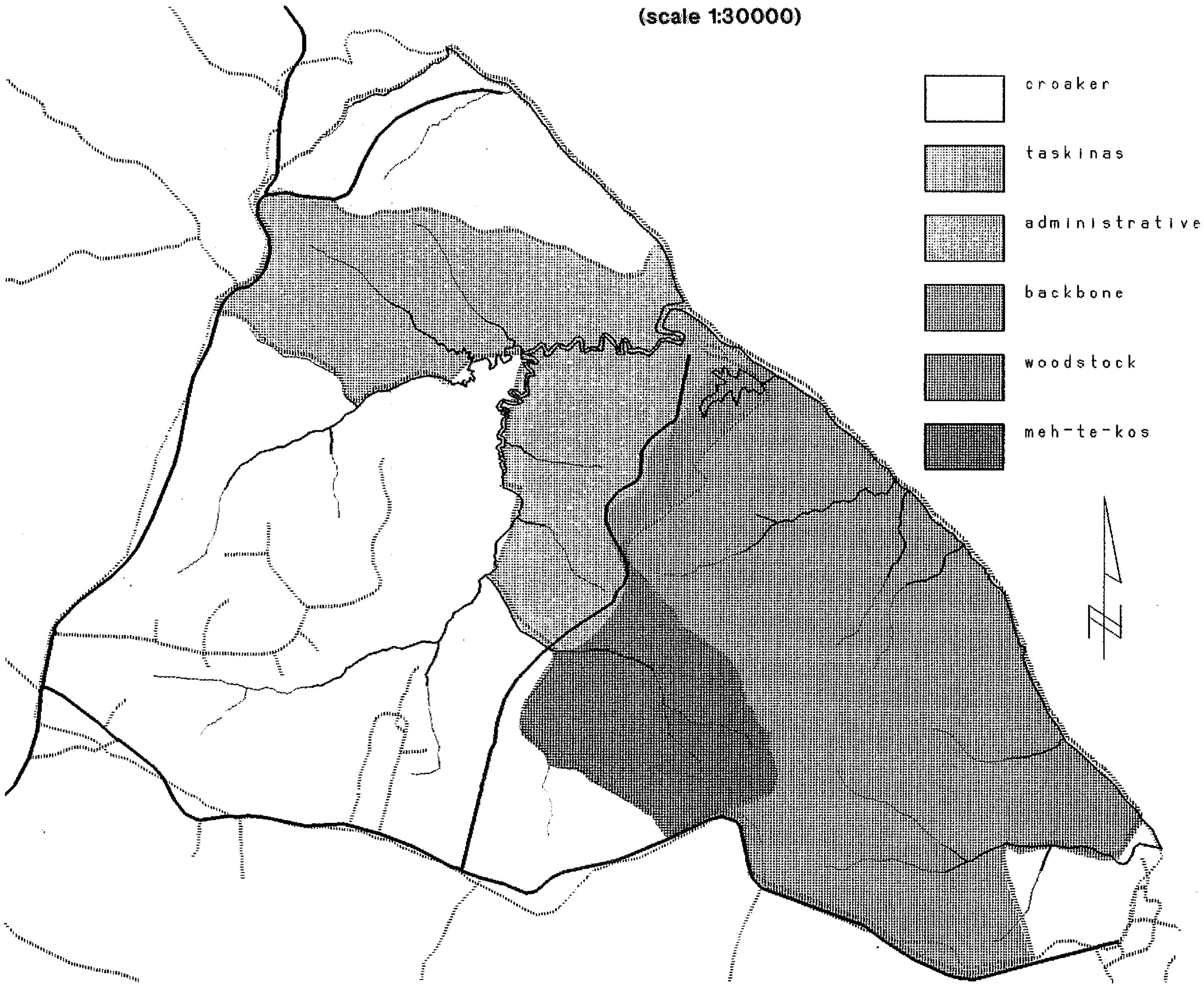
Since the Botanical Unit overlapped with several of the other units' boundaries which were based on use prior to receipt of this information, it was decided that the Botanical Unit would be created as a separate layer. Two different methods were used to create this coverage: manual transfer and optically scanned digitization.

<u>Manual transfer</u>:

MGTBOUNIT was created, edited and built using methods described in earlier sections. Each coverage was edited and built. Unit names (AREANAME) and Codes (LU-CODE) were added to each coverage's Polygon Attribute Table file.

Figure 5 : Resource Management Units

(scale 1:30000)



Optically scanned:

The first step in bringing these boundaries into the GIS was to trace them from the original map onto acetate using a technical pen and india ink. In addition to the boundaries, geographic control points were traced onto the acetate. These included tic-marks and identifiable landmarks (e.g. road intersections).

The acetates were then scanned [PREPARE \ SCAN] into MIPS (Map and Image Processing System, vers. 2.90) as binary (black or white) rasters with the raster displayed as white lines on a black background. After some initial experimentation, the selected settings for scanning were 150 dots per inch and a threshold of 125.

The resulting rasters were then edited [PREPARE \ RASTER \ EDIT-RASTER] to fill in pen skips and to remove specks, stray pen marks and other undesirable marks. The two primary tools used for this purpose were PAINT and SMARTFILL. SMARTFILL works in two steps, first delineating the polygon surrounding the cursor and then FLOODFILLing it. If no polygon is found in the first step nothing can be filled. For this reason, SMARTFILL is of particular utility for ensuring that polygons are indeed closed. The next step in editing was to THIN the lines to a single pixel in width. After the lines had been THINned, the raster was checked for extra line segments and other errors, which were corrected with PAINT.

The raster was then converted to a vector using the automatic raster-to-vector conversion facility in MIPS [PREPARE \ RASTER->VECTOR \ AUTO-LINES]. Polygons and lines were then labeled [PREPARE \ VECTOR \ EDIT \ CLASS+TEXT]. The resulting vector was then registered to the original 7.5' topographic map with a GITICO digitizer [PREPARE \ VECTOR \ MAP PROJ \ MANUAL \ CREATE+EDIT]. The registration of the vector was then converted from latitude\longitude to UTM zone 18 [PREPARE \ VECTOR \ MAP PROJ \ CHANGE PROJ]. Finally, the vector was exported [PREPARE \ EXPORT \ VECTOR \ EXPORT VECTOR] in DLG optional format, in MOSS format, and in ARC-generate format. While all three of these formats were workable, DLG was easiest to bring into ARC. (An initial export in latitude \ longitude proved to be more difficult to work with.)

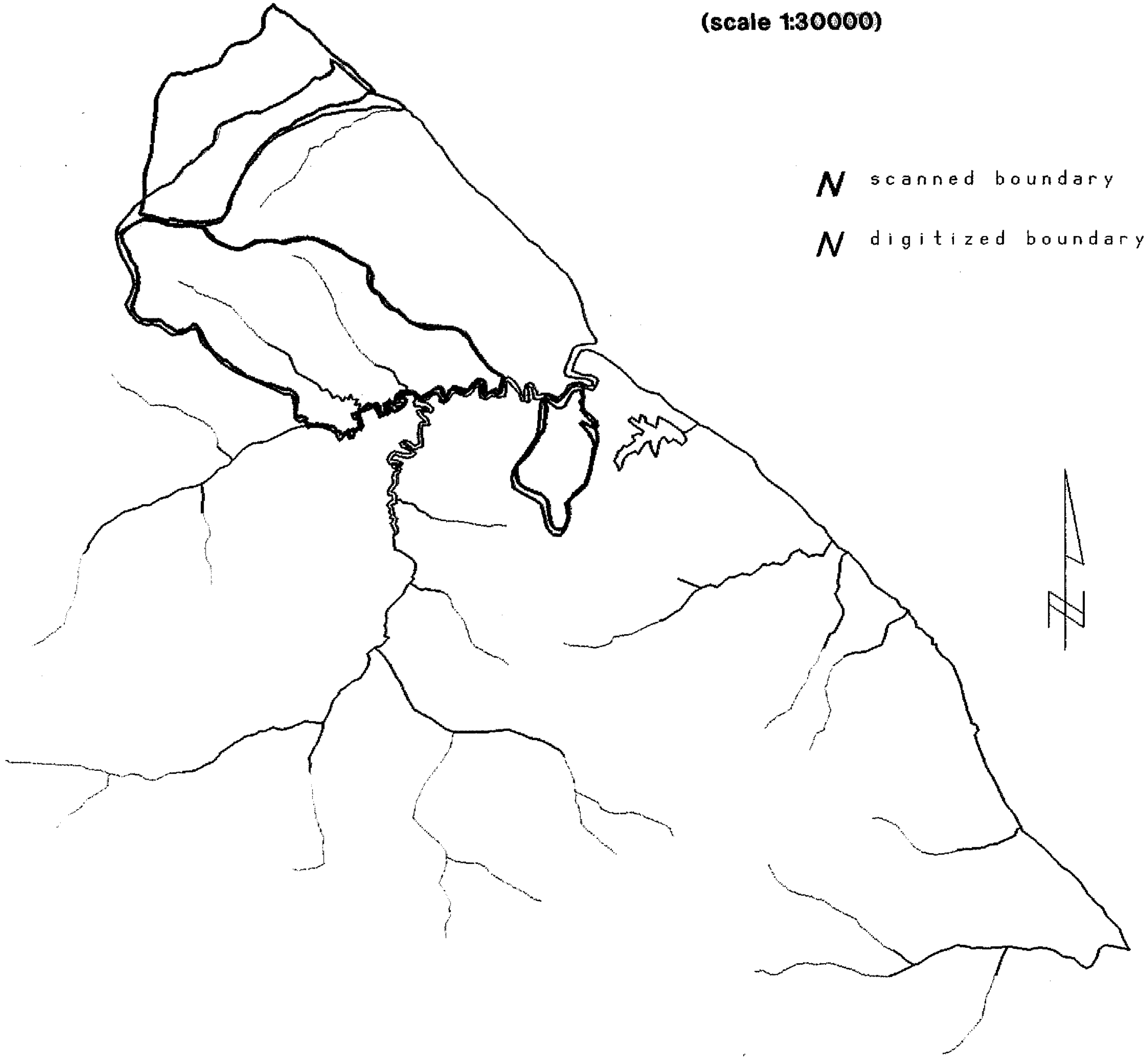
In Figure 6, it is apparent that, while the configuration of both coverages is similar, there is a significant error in location. Differences in line locations vary as much as 700 feet, but appear to average 100 feet. Unfortunately, at this time it is not possible to determine which of the two methods is most accurate due to the methods of transferring and digitizing datalayers.

The digitized coverage was hand-transferred from color photocopy to a 7.5' topoquad, which may have introduced some operator error. Likewise, the scanned image was taken from an acetate sheet on which the original image had been traced, again potentially introducing error. Differences in scale (the manually transferred coverage was digitized at 1:24,000 while the scanned image was fitted to a scanned 7.5' topoquad) - may have introduced additional error.

Since scanning appears to be the least time-consuming of the two methods it would be advantageous if this method is accurate. We are hoping to follow up on this possibility

Figure 6 : Comparison of Scanned versus Digitized Datalayer Development

(scale 1:30000)



sometime in 1992, using an acetate sheet original overlaid on a map from which both scanned and manually digitized coverages can be developed.

The final coverages are MGTUNIT.CL (for the general Management Units) and MGTBOUNIT (for the Botanical Mgt Units). (See Appendix A.)

6) National Wetlands Inventory - Rectification with Base Map Datalayers.

Upon reviewing base map and product overlays, it was noted that significant discrepancies existed between digitized streams data and NWI wetland polygon and line locations. (See Figure 7.) Intuitively, NWI line data and streams should conjointly occur. Firsthand knowledge of the area supports the assumption that these linear features are closely-associated and should overlap in most instances. The error is due to problems associated with aerial photo-interpretation, parallax and image displacement. The closer an object is to the edge of the frame (i.e., the farther it is from photo center) the more displaced is its apparent location. To correct this problem, it was decided that the digitized 'streams' data would be accepted as the controlling data, and an effort was made to edit the NWI coverage.

The 'Streams' coverage was copied into a 'Wetland' coverage. The new Wetland coverage was edited with the original Streams coverage used as an overlay. Some lines were deleted, copied, split and merged with shorter lines, and edited to line up with the streams. (Figure 7.)

During the process, it was also noted that polygon locations appeared inconsistent as well (based on an examination of stream locations and marsh symbols on the USGS topoquad). However, editing this coverage would not be easy, since moving the polygons would be arbitrary (based on what appears to be the proper location) and possibly influenced by polygon morphology, image displacement and the number and placement of nodes within polygons. To approximate locations of polygons without substantive documentation to support relocation was not believed to be useful and the editing was not completed.

If a vector image could be created (overlaid) on a raster image, using MIPS, it may be possible to move this data to appropriate locations. Alternatively, it may be worthwhile to investigate the potential for overlaying and rubber-sheeting vector data on a raster image and then converting that file into Arc/Info format and projections. (This would require rubber-sheeting data twice and may unacceptably distort the data.)

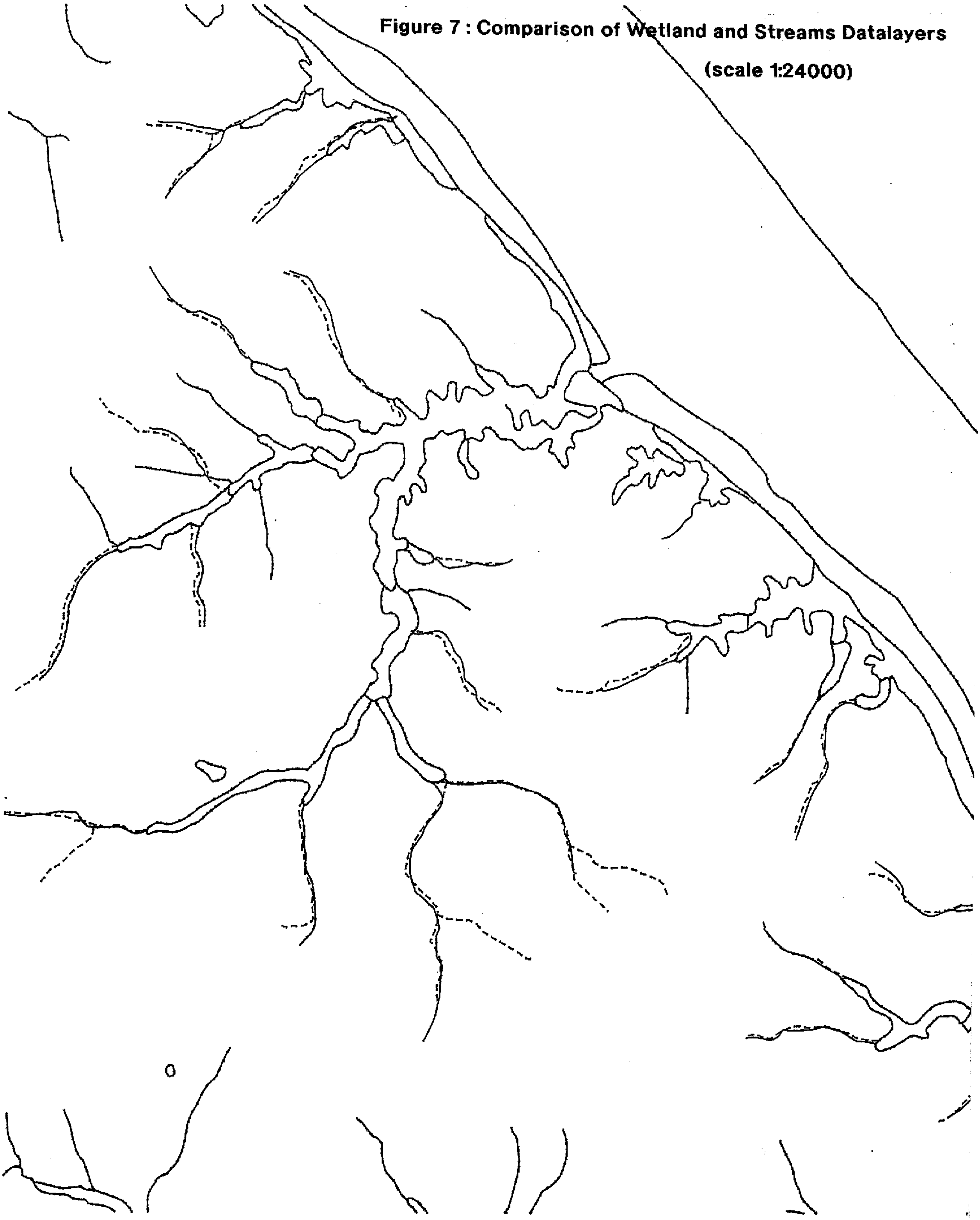


Figure 7 : Comparison of Wetland and Streams Datalayers

(scale 1:24000)

III. SUMMARY OF PROBLEMS

-DLG data does not include all pertinent information for individual datalayers. Specifically, some stream segments were absent. In addition, DLG road data may not coincide with roads drawn on topoquads which may introduce error.

For site-specific management, it may be best to choose a smaller scale from which to work and digitize all information in at that scale, or a smaller one.

-Wetlands datalayer does not agree with digitized stream configurations or empirical knowledge of on-site associations.

Wetlands informations is a useful tool in properly managing individual sites. However, there are limitations in using existing datalayers as more than an indication of potential wetlands locations/configuration. When using this information, uncorrected, care must be taken to insure that it is not overlain on contradictory information or, if it is that the limitations of its use are acknowledged.

-Forest datalayer development is difficult using current mapping methods.

Initially, the problem has been transferring information from a crudely drawn forest inventory map to the base map. This is compounded by the distinction of several different forest community types. Several alternatives are possible here which may be of value.

-Require that all future forest inventories be mapped on USGS topoquads. Again, some operator error is associated with this, but individuals mapping this information should be better suited to transfer field information to a standard base map than individuals operating the GIS.

-Map forest type using an acetate overlay on one photograph, and scanned imagery to fit the image to a USGS.

- Ground-truth existing forest cover on USGS topoquad, noting changes; use aerial photos to distinguish major types (i.e. evergreen, deciduous). This would produce good scale maps, but the level of information would decrease substantially.

IV. ESTIMATE OF STAFFING TIME TO DEVELOP GIS DATA LAYERS FOR COASTAL PARKS

Part of the purpose of the grant was to estimate the amount of time and money necessary to expand this process to the other coastal parks (and subsequently, other regions of the state). Time commitments for each datalayer were estimated, since in many instances several datalayers would be in different stages of development at the same time, making direct measurement of time expenditures difficult. Time estimates, source and coverage scales and method of digitization are found in Appendix B.

Approximately twenty-six full days were necessary to digitize coverages, transfer existing coverages into the appropriate Arc/Info files and develop databases for those coverages

(Appendix B). This was based on one person working full time on just GIS development. Coordination of the work load and trouble-shooting would take additional time. Therefore, a more realistic estimate may be 35 days.

Given these estimates, approximately seven parks (the remainder of the coastal parks) could be completed in a year's time, with one employee dedicated to work on the GIS. If a P-14 were used, then approximately 1500 hours a year could be devoted to GIS development. This would allow, five parks to be added to the GIS annually.

Assuming a salary of $10.00 per hour, this translates to a financial commitment of $15,000 to 26,000 on an annual basis, for salary and fringe benefits. Costs of overhead and supplies would add some to this cost.

Expanding this to all parks, the Department would need to make an annual commitment, or seek outside funding, at this level for three to four years to complete the entry of data on all parks into the Arc/Info system, using full time employees. If P-14 staff were used, then the estimate increases to over five years.

V. UTILITY OF GIS TO THE DEPARTMENT

The utility of GIS to this department is apparent in that the Council on Information Management has recognized DCR as "major provider and user of" GIS technology. The Department currently makes use of GIS in three modes: ARC/INFO on a minicomputer, ARC/INFO on personal computers, and VirGIS. Each of these access a wealth of information that is currently in electronic form.

Although there is a great deal of information available, current GIS capabilities within DCR are used to the greatest extent by the Division of Soil and Water Conservation, with four people working on a regular basis with Arc/Info. These people have current commitments that essentially limit their ability to work on departmental efforts. Other people have been trained and there are PC Arc/Info and graphic terminal connections to the Prime within the Department. However, existing job responsibilities limit the ability of others to devote the necessary blocks of time to become proficient with the software package. Therefore, GIS must be integrated into the existing work programs and used as a mechanism to accomplish these work efforts, to realize its potential as a planning tool.

The coverages that have been added to the GIS in this grant (watershed, land classification, management units, forest type) will make site management easier in that comparisons can be made in the office (and then substantiated in the field). Additionally, information on forest cover types, open areas, land classification, streams and wetlands can be used as a base upon which to build other data layers. Complex relationships between physical and biological variables may be more readily apparent when graphically displayed, and subsequently managed more effectively. For example, with the addition of soils information, rare species or forest types might be correlated to this and other information, serving to pinpoint other potential locations within the park. Likewise, soils and land use information may lead to changes in existing use. Overlaying management units on these coverages can fine-tune management prescriptions or objectives to effectively protect sensitive areas within the park.

Ultimately, GIS will be used to evaluate existing land classifications and assist in site assessments for future development and habitat management.

Suggestions for Future Work

In working with biological and physical features in GIS, a knowledge of biological sciences, planning and cartographic techniques would be especially valuable in the development of coverages in the future and should be a consideration in employment of GIS technicians on future projects. Although an understanding of GIS is imperative, it is felt that an appreciation of the inter-relationship between the individual coverages being developed is almost as important and should not be overlooked.

This grant effort showed the use of a MIPS - Arc/Info transfer to be a potentially valuable and powerful tool in the development of new coverages. The amount of time devoted to scanning is about half that of manual digitization of the coverage and may prove to be valuable in transferring coverages to Arc/Info. Current concerns with accuracy need to be more fully evaluated at this point.

Efforts should be made to investigate the use of MIPS' ability to overlay raster and vector information to create rectified or "photo-corrected" vector sets.

Satellite imagery is currently available in a rectified form which may serve as a small scale base from which rectified products (such as wetlands maps) may be developed. This should be investigated, along with the possibility of using satellite imagery and aerial photography to create higher resolution basemaps.

Work should be done to develop an appropriate scale for all base map development. Although the 7.5' USGS topoquad is relatively standard, for site-specific work this scale may be too small. Additionally, the use of same-scale maps would be required in future work and digitization should be done at a larger scale than the one intended for standard use to improve the appearance of the final product. (For example, digitization should be done at 1:12000) if the standard scale is 1:24000.)

Coverages should be developed for areas outside park boundaries within the watershed which include land use, soils, water features, etc.

VI. SUMMARY

The development of GIS datalayers for the Division of State Parks is possible using information that is available as part of its resource management planning process. A great deal of information already exists in Arc/Info (soils, roads, wetlands) that is useful in managing our resources. Some of the information to be digitized may need to be modified or evaluated more closely than is currently done (e.g. forestry datalayers) to insure that the level of accuracy is maintained.

Scanning technology appears to be a very useful tool which may make the process flow more quickly and one of the things learned from this grant is that this should be more effectively evaluated in the future.

As mentioned earlier, all of this work will require a substantial commitment of time and resources if information on all parks is to be entered into Arc/Info within a reasonable time period (i.e. three to four years) and may best be accomplished using staff dedicated to GIS development.

Literature Cited

Gaston, Dennis. 1989. Forest Inventory of York River State Park. Virginia Department of Forestry, unpublished report.

DCR. 1990. Land Classification System: Land Classification Committee unpublished report.

DCR. 1991. York River State Park Resource Management Plan, Division of State Parks, unpublished report.

Appendix A: Plotfile Coverages

Plotfile Coverage:	Plotfile:
1. Base Map YRH20 SHED.UTM18 ROADS YRSP/STRMEPH.UTM18 YRSP/STRMPER.UTM18 PARKBND	#BASE
2. Streams YRSP/STRMEPH.UTM18 YRSP/STRMPER.UTM18	#STREAM
3. Forest	
MGMT.UTM	#FOREST
FORMGT2.UTM18	#FORMGT #FORMGMT
4. Watershed Boundary	
YRH20 SHED.UTM18	#WATERSHED
5. National Wetlands Inventory W095 E	
WETLAND.UTM18	#WETLAND
6. Management Units	
MGTUNIT.CL	#MGTUNIT #MANAGEMENT
MGTBOUNIT	#MGTBOUNIT #BOTANICAL
7. Historical/Cultural Sites	
HISTORIC	#HISTORIC
8. Land Classification	
LNDCLS.CL	#LNDCLS #LAND

Appendix B: Time Estimates for Data Entry

Data Layer	SCALE Source Scale	 Digitized Product	Method	Time
Roads	1:100000	1:100000	DT[1]	2 hours
Wetlands	1:24000	1:24000	DT	2 hours
Streams	1:24000	1:24000	DD	2 days
Park Boundary	1:24000	1:24000	DD	2 hours
Watershed	1:24000	1:24000	DD	2 days
Forest Cover	1:25200	1:24000	PE/MT	10 days
Land Classification	1:12000	1:24000	PR/MT	5 days
Management Units	1:12000	1:24000	PR/MT	5 days
Botanical Management Units	-	1:24000	MT[2] OS	1 day 4 hours

[1] DT - Digital transfer
DD - Direct digitization from topoquad
PE - Photo-enlargement
PR - Photo-reduction
MT - Manual transfer of locally fit source and digitized
OS - Optically scanned and fit to a scanned image of the Gressit topoquad converted to DLG, Arc Generate and MOSS formats using MIPS

[2] This data layer was an enlarged color photocopy of the topoquad at an unknown scale which allowed manual transfer based on contour lines. Photo-reduction was not used.